

JA999075

日 本 国 特 許 庁

PATENT OFFICE
JAPANESE GOVERNMENT

JCS31 U.S. PRO
09/741715
12/20/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出 願 年 月 日

Date of Application:

1999年12月22日

出 願 番 号

Application Number:

平成11年特許願第364627号

願 人

Applicant(s):

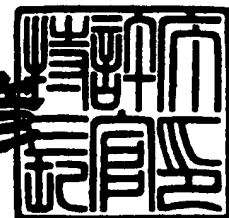
インターナショナル・ビジネス・マシーンズ・コーポレーシ
ョン

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 6月29日

特許庁長官
Commissioner,
Patent Office

近 藤 隆 彦



出証番号 出証特2000-3051302

【書類名】 特許願

【整理番号】 JA999075

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/00

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 東京基礎研究所内

【氏名】 立花 隆輝

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 東京基礎研究所内

【氏名】 清水 周一

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 東京基礎研究所内

【氏名】 小林 誠士

【特許出願人】

【識別番号】 390009531

【住所又は居所】 アメリカ合衆国 1 0 5 0 4、ニューヨーク州アーモンク (番地なし)

【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】 100086243

【弁理士】

【氏名又は名称】 坂口 博

【連絡先】 0 4 6 2 - 7 3 - 3 3 1 8、3 3 2 5、3 4 3 1

【選任した代理人】

【識別番号】 100091568

【弁理士】

【氏名又は名称】 市位 嘉宏

【手数料の表示】

【予納台帳番号】 024154

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9304391

【包括委任状番号】 9304392

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 圧縮オーディオデータへの電子透かし方法およびそのシステム

【特許請求の範囲】

【請求項 1】

圧縮オーディオデータに付加情報を埋め込むシステムであって、

- (1) 圧縮オーディオデータからMDC T係数を復元する手段と、
 - (2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、
 - (3) 求めた前記周波数成分に対して、付加情報を周波数空間で埋め込む手段と、
 - (4) 前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する手段と、
 - (5) 付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する手段、
- を有する、付加情報埋め込みシステム。

【請求項 2】

圧縮オーディオデータに埋め込まれた付加情報を更新するシステムであって、

- (1) 圧縮オーディオデータからMDC T係数を復元する手段と、
 - (2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、
 - (3) 求めた前記周波数成分から、付加情報を検出する手段と、
 - (3-1) 前記周波数成分の前記付加情報を必要に応じて変更する手段と、
 - (4) 前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する手段と、
 - (5) 付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する手段、
- を有する、付加情報埋更新システム。

【請求項 3】

圧縮オーディオデータに埋め込まれた付加情報を検出するシステムであって、

- (1) 圧縮オーディオデータからMDC T係数を復元する手段と、
 - (2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、
 - (3) 求めた前記周波数成分から、付加情報を検出する手段と、
- を有する、付加情報検出システム。

【請求項 4】

前記オーディオデータの周波数成分を求める手段(2)が、MDC T係数と周波数成分の対応関係を含む既定のテーブルを用いて、周波数成分を求める手段である、請求項 1 乃至 3 の何れかに記載のシステム。

【請求項 5】

前記周波数成分をMDC T係数に変換する手段(4)が、MDC T係数と周波数成分の対応関係を含む既定のテーブルを用いて、MDC T係数に変換する手段である、請求項 1 乃至 2 の何れかに記載のシステム。

【請求項 6】

前記付加情報を周波数空間で埋め込む手段(3)が、1ビットを埋め込む領域を時間領域で分割し、その各部分について信号レベルを計算し、各周波数ごとに最弱な信号レベルにあわせて、付加情報を周波数空間で埋め込む手段である、請求項 1 記載のシステム。

【請求項 7】

MDC T係数と周波数成分の対応関係を含むテーブルを作成する方法であって、圧縮データの圧縮に用いられた、少なくとも1つの窓関数及び窓長について、

- (1) 時間軸上の波形に対してフーリエ変換を行うときの基底を作成する段階と、
 - (2) 前記基底を用いて生成される波形に、対応する窓関数を乗じる段階と、
 - (3) 前記窓関数を乗じた結果にMDC Tを行い、MDC T係数を算出する段階と、
 - (4) 前記基底と前記MDC T係数とを対応付ける段階と、
- を有する、テーブル作成方法。

【請求項 8】

前記対応する窓関数を乗じる段階（２）が、前記基底の周期性を利用することにより、冗長な、周波数成分とMDC T係数の対応関係を生成しない、請求項 7 記載のテーブル作成方法。

【請求項 9】

前記対応する窓関数を乗じる段階（２）が、前記基底を幾つかの部分に分割し、それぞれに対応する窓関数を乗じることにより、冗長な、周波数成分とMDC T係数の対応関係を成しない、請求項 7 記載のテーブル作成方法。

【請求項 1 0】

圧縮オーディオデータに付加情報を埋め込む方法であって、

- （１）圧縮オーディオデータからMDC T係数を復元する段階と、
 - （２）復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める段階と、
 - （３）求めれた前記周波数成分に対して、付加情報を周波数空間で埋め込む段階と、
 - （４）前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する段階と、
 - （５）付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する段階、
- を有する、付加情報埋め込み方法。

【請求項 1 1】

圧縮オーディオデータに埋め込まれた付加情報を更新する方法であって、

- （１）圧縮オーディオデータからMDC T係数を復元する段階と、
- （２）復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める段階と、
- （３）求めれた前記周波数成分から、付加情報を検出する段階と、
- （３－１）前記周波数成分の前記付加情報を必要に応じて変更する段階と、
- （４）前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する段階と、
- （５）付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作

成する段階、

を有する、付加情報埋更新方法。

【請求項 1 2】

圧縮オーディオデータに埋め込まれた付加情報を検出する方法であって、

(1) 圧縮オーディオデータからMDCT係数を復元する段階と、

(2) 復元された前記MDCT係数を用いて、オーディオデータの周波数成分を求める段階と、

(3) 求めた前記周波数成分から、付加情報を検出する段階と、

を有する、付加情報検出方法。

【請求項 1 3】

前記オーディオデータの周波数成分を求める段階(2)が、MDCT係数と周波数成分の対応関係を含む既定のテーブルを用いて、周波数成分を求める段階である、請求項 1 0 乃至 1 2 の何れかに記載の方法。

【請求項 1 4】

前記周波数成分をMDCT係数に変換する段階(4)が、MDCT係数と周波数成分の対応関係を含む既定のテーブルを用いて、MDCT係数に変換する段階である、請求項 1 0 乃至 1 1 の何れかに記載の方法。

【請求項 1 5】

請求項 7 ～ 9 の何れかに記載のテーブル作成方法を実行するためのプログラムを記憶した、コンピュータ読取可能なプログラム記憶媒体。

【請求項 1 6】

請求項 1 0 に記載の付加情報を埋め込む方法を実行するためのプログラムを記憶した、コンピュータ読取可能なプログラム記憶媒体。

【請求項 1 7】

請求項 1 1 に記載の付加情報を更新する方法を実行するためのプログラムを記憶した、コンピュータ読取可能なプログラム記憶媒体。

【請求項 1 8】

請求項 1 2 に記載の付加情報を検出する方法を実行するためのプログラムを記憶した、コンピュータ読取可能なプログラム記憶媒体。

【請求項 1 9】

付加情報を埋め込む対象となる圧縮オーディオデータに前記付加情報を埋め込む情報埋め込み装置と、前記圧縮オーディオデータに付加情報が埋め込まれたか否かを検出する検出装置とを有する電子透かし装置であって、

前記情報埋め込み装置は、

- (1) 圧縮オーディオデータからMDC T係数を復元する手段と、
- (2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、
- (3) 求めた前記周波数成分に対して、付加情報を周波数空間で埋め込む手段と、
- (4) 前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する手段と、
- (5) 付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する手段、

を有し

前記検出装置は、

- (6) 圧縮オーディオデータからMDC T係数を復元する手段と、
- (7) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、
- (8) 求めた前記周波数成分から、付加情報を検出する段階と、

を有する

電子透かし装置。

【発明の詳細な説明】

【0 0 0 1】

【産業上の利用分野】

本発明は、圧縮されたデジタルオーディオデータに対して著作権情報等の付加情報の埋め込み、検出、更新を行う方法とそのシステムに関し、特に周波数空間での電子透かし技術と等価の操作を圧縮されたオーディオデータに対して適用可能とする技術に関する発明である。

【 0 0 0 2 】

【従来の技術】

オーディオデータへの電子透かし技術には、Spread Spectrum 法、ポリフェーズフィルタを用いる方法、周波数空間へと変換した後に埋め込みを行う方法などがある。周波数空間で埋め込み・検出を行う方法には、聴覚心理モデルの適用が容易で、高音質を実現しやすいこと、変換やノイズに対する耐性が強いという利点がある。しかし従来の音声電子透かし技術の対象は、圧縮処理を施されていない状態にあるデジタル・オーディオデータに限定されていた。オーディオデータのインターネット配信においては、通信容量の制限からオーディオデータを音声圧縮して利用者へと配信するのが通常であって、従来の電子透かし技術を適用するには圧縮状態を解凍し、埋め込みを行い、再び再圧縮を行う必要があった。そして、高音質と高圧縮効率を同時に実現している先進的な音声圧縮技術であればあるほど、この一連の操作に必要な計算時間は長くならざるを得ない。オーディオデータを聴くことができるまでにかかる時間は利用者の購買意欲に大きな影響を及ぼす。従ってオーディオデータを圧縮した状態のままで、付加情報の埋め込み、変更、検出を行うことが要望される。しかしながら、圧縮状態にあるデジタルオーディオデータに付加情報を直接埋め込み、これを変更、検出する方法は知られていない。

【 0 0 0 3 】

【発明が解決しようとする課題】

従って、本発明が解決しようとする課題は、上記問題点に鑑み発明されたものであり、圧縮状態にあるデジタルオーディオデータ内の情報を直接操作する方法およびシステムを提供することである。

また別の課題は、圧縮状態にあるオーディオデータに付加情報を埋め込む方法およびシステムを提供することである。

また別の課題は、デジタルオーディオデータに少ないメモリ容量で付加情報を埋め込む方法およびシステムを提供することである。

また別の課題は、デジタルオーディオデータに埋め込む付加情報を、最小にして埋め込む方法およびシステムを提供することである。

また別の課題は、圧縮されたデジタルオーディオデータにすでに埋め込まれている付加情報を、圧縮された状態で検出する方法およびそのシステムを提供することである。

また別の課題は、圧縮されたデジタルオーディオデータにすでに埋め込まれている付加情報を、圧縮された状態で変更する方法およびシステムを提供することである。

【0004】

【課題を解決するための手段】

〔付加情報埋め込みシステム〕

上記課題を解決するために、本発明の圧縮オーディオデータに付加情報を埋め込むシステムは、

(1) 圧縮オーディオデータからMDCT (Modified Discrete Cosine Transform) 係数を復元する手段と、

(2) 復元された前記MDCT係数を用いて、オーディオデータの周波数成分を求める手段と、

(3) 求めた前記周波数成分に対して、付加情報を周波数空間で埋め込む手段と、

(4) 前記付加情報の埋め込まれた周波数成分をMDCT係数に変換する手段と、

(5) 付加情報の埋め込まれた前記MDCT係数から圧縮オーディオデータを作成する手段、

を有する。

【0005】

〔付加情報更新システム〕

また、本発明の圧縮オーディオデータに埋め込まれた付加情報を更新するシステムは、

(1) 圧縮オーディオデータからMDCT係数を復元する手段と、

(2) 復元された前記MDCT係数を用いて、オーディオデータの周波数成分を求める手段と、

(3) 求めれた前記周波数成分から、付加情報を検出する手段と、

(3-1) 前記周波数成分の前記付加情報を必要に応じて変更する手段と、

(4) 前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する手段と

(5) 付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する手段、

を有する。

【0006】

〔付加情報検出システム〕

また、本発明の圧縮オーディオデータに埋め込まれた付加情報を検出するシステムは、

(1) 圧縮オーディオデータからMDC T係数を復元する手段と、

(2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、

(3) 求めれた前記周波数成分から、付加情報を検出する手段と、

を有する。

【0007】

好ましくは、前記オーディオデータの周波数成分を求める手段(2)は、MDC T係数と周波数成分の対応関係を含む既定のテーブルを用いて、周波数成分を求める。

【0008】

好ましくは、前記周波数成分をMDC T係数に変換する手段(4)は、MDC T係数と周波数成分の対応関係を含む既定のテーブルを用いて、MDC T係数に変換する。

【請求項20】

好ましくは、前記付加情報を周波数空間で埋め込む手段(3)は、1ビットを埋め込む領域を時間領域で分割し、その各部分について信号レベルを計算し、各周波数ごとに最弱な信号レベルにあわせて、付加情報を周波数空間で埋め込む。

【0009】

〔対応テーブル作成方法〕

また、本発明のMDC T係数と周波数成分の対応関係を含むテーブルを作成する方法は、圧縮データの圧縮に用いられた、少なくとも1つの窓関数及び窓長について、

- (1) 時間軸上の波形に対してフーリエ変換を行うときの基底を作成する段階と、
 - (2) 前記基底を用いて生成される波形に、対応する窓関数を乗じる段階と、
 - (3) 前記窓関数を乗じた結果にMDC Tを行い、MDC T係数を算出する段階と、
 - (4) 前記基底と前記MDC T係数とを対応付ける段階と、
- を有する。なお基底の例としては正弦波、余弦波などがあげられる。

【0 0 1 0】

〔付加情報埋め込みシステムの作用〕

本発明の圧縮オーディオデータに付加情報を埋め込むシステムは、まず、圧縮されたデジタルオーディオデータから圧縮されていたMDC T係数を復元する。あらかじめ計算してテーブルに記憶しておいたMDC T係数列を用いて、オーディオデータの周波数成分を求める。これに対し周波数空間における付加情報の埋め込み方法を用い、埋め込み周波数信号を計算する。求めた埋め込み周波数信号を前記テーブルを用いて再びMDC T係数へと変換し、オーディオデータのMDC T係数に加算し、これを新たなオーディオデータのMDC T係数とする。このMDC T係数を再び圧縮を施し、埋め込み後のデジタルオーディオデータとする。

【0 0 1 1】

さらに、本発明の最小埋め込み方法は、1ビットを埋め込むフレームを時間領域で分割し、その各部分について信号レベルを計算し、各周波数ごとに最弱な信号レベルにあわせて埋め込み信号の上限を計算するように構成する。

【0 0 1 2】

〔対応テーブルの作用〕

本発明のMDC T係数と周波数成分の対応テーブルは、フーリエ変換の各基底がMDC T係数にどのように表現されるかをフレーム長（窓関数、窓長）に応じ

てあらかじめ計算したテーブルを作成する。これにより圧縮状態にあるオーディオデータの直接操作を行うことができる。

【0013】

本発明の対応テーブルに要求されるメモリサイズの縮小手段は、正弦波、余弦波などの基底の周期性を利用することにより、冗長な情報を記憶しないようにする。またはフーリエ変換の各基底をそのままMDCTした結果をテーブルに保存するのではなく、各基底を幾つかの部分に分割しそれぞれに対応するMDCT係数を保存することでテーブルの記憶に必要なメモリサイズを縮小する。

【0014】

〔付加情報検出システムの作用〕

本発明の圧縮オーディオデータに埋め込まれた付加情報を検出するシステムは、符号化されていたMDCT係数を復元し、埋め込みシステムと同様のテーブルを用いて、周波数空間での検出と等価な操作を行いビット情報や符号信号を検出する。

【0015】

〔付加情報更新システムの作用〕

本発明の圧縮オーディオデータに埋め込まれた付加情報を更新するシステムは、符号化されていたMDCT係数を復元し、検出システムと同じ方法を用いてこのMDCT係数から埋め込まれた信号の検出を行う。その信号が十分な強度を持っていない場合、あるいは埋め込む信号とは異なる信号が検出され更新を行う必要がある場合のみ、埋め込みシステムと同じ方法を用いてMDCT係数に埋め込みを行う。得られた新しいMDCT係数を再び符号化し更新後のデジタルオーディオデータとする。

【0016】

【発明の実施の形態】

まず本発明の実施の形態を説明する前に語句の定義を行う。

「音声圧縮技術」

本発明が対象とする圧縮データは、主として音声、音楽、効果音など音全般を電子的にデータ化し、これを圧縮したものである。音の圧縮技術は MPEG1、MPEG2

、MP3 などとして知られている。明細書中では、このような圧縮技術を総合して音声圧縮技術と呼ぶ。また音全般を簡潔に音声もしくはオーディオとして記載する。

・圧縮状態

対象の音声圧縮技術によって音声データが、音声の劣化を最低限度にとどめつつデータ量を減らされている状態を呼ぶ。

・非圧縮状態

WAVE ファイルや AIFF ファイルなど、音声の波形が加工なく記述されている状態を指す。

・圧縮状態をほどく

音声データを「圧縮状態から非圧縮状態へと変換する」ことを指す。「非圧縮状態へ移す」も同義である。

・MDC T変換 (Modified Discrete Cosine Transform)

【数 1】

$$M_k = \sum_{n=0}^{N-1} X_n \cos \left\{ \frac{2\pi}{N} \left(n + \frac{N}{4} + \frac{1}{2} \right) \left(k + \frac{1}{2} \right) \right\}$$

X_n は時間軸上のサンプル値であり n は時間軸方向のインデクスである。

M_k がMDC T係数であり、 k は0から $(N/2)-1$ の整数で周波数を示すインデクスである。

この操作によって時間軸上の系列 $X_0 \sim X_{(N-1)}$ を周波数軸上の系列 $M_0 \sim M_{((N/2)-1)}$ に変換するのがMDC T変換である。MDC T係数も一種の周波数成分を表しているが、本明細書中では「周波数成分」という言葉ではDFT変換の結果として得られる係数のことを指す。

・DFT変換 (離散フーリエ変換, Discrete Cosine Transform)

【数 2】

$$R_k = \sum_0^{N-1} X_n \cos \left\{ \frac{2\pi}{N} kn \right\}$$

$$I_k = - \sum_0^{N-1} X_n \sin \left\{ \frac{2\pi}{N} kn \right\}$$

X_n は時間軸上のサンプル値であり n は時間軸方向のインデクスである。

R_k が実数成分（余弦波成分）、 I_k が虚数成分（正弦波成分）であり、 k は0から $(N/2)-1$ の整数で周波数を示すインデクスである。この操作によって時間軸上の系列 $X_0 \sim X_{(N-1)}$ を周波数軸上の系列 $R_0 \sim R_{((N/2)-1)}$ および $I_0 \sim I_{((N/2)-1)}$ に変換するのが離散フーリエ変換である。本明細書中では「周波数成分」と呼ぶのはこの R_k と I_k の両方の系列の総称である。

・窓関数

MDC Tを行う前にサンプルに乗算される関数である。一般にサイン関数やカイザー関数などが使われる。

【0 0 1 7】

・窓長

音声データの特性に応じて、データに乗じる窓関数の形状やその長さを指し、MDC Tを行う際に幾つのサンプルに対してMDC Tを行うかを表す値である。

【0 0 1 8】

図 1 に圧縮オーディオデータに付加情報を直接埋め込む装置のブロック図を示す。ブロック 1 1 0 は圧縮オーディオデータを入力として、MDC T 係数列を復元するブロックである。ブロック 1 2 0 は、ブロック 1 2 0 で復元された MDC T 係数を用いて、オーディオデータの周波数成分を求めるブロックである。ブロック 1 3 0 は、ブロック 1 2 0 で求めれた周波数成分に対して、付加情報を周波数空間で埋め込むブロックである。ブロック 1 4 0 は、ブロック 1 3 0 で付加情報の埋め込まれた周波数成分を MDC T 係数に変換するブロックである。そして最

後にブロック 1 5 0 で、ブロック 1 4 0 で変換された M D C T 係数から圧縮オーディオデータを作成する。

【 0 0 1 9 】

上記ブロック 1 2 0 とブロック 1 3 0 では、M D C T 係数・周波数の対応テーブルを用いて変換を高速に行う。本発明ではフーリエ変換の各基底が M D C T 空間内においてどのように表現されるかをテーブルにあらかじめ保存しておき、それを埋め込み・検出・更新の各システムに利用する。以下に、M D C T 係数・周波数の対応テーブルとその作成方法、圧縮されたオーディオ・データに対する埋め込みシステム、検出システム、更新システム、そして関連するその他の方法を説明する。

【 0 0 2 0 】

[M D C T 係数・周波数の対応テーブル]

埋め込み時の演算に聴覚心理モデルを利用するためにはオーディオデータを周波数空間へと変換する必要があるが、M D C T 係数として表現されたオーディオデータを時間軸上へと逆変換しフーリエ変換を行うことで求めるには多大な計算時間が必要となる。そこで、M D C T 係数と周波数成分の直接的な対応関係を知る必要がある。

【 0 0 2 1 】

もし一定のサンプル数に対して窓関数なしで M D C T をほどこしてオーディオデータが圧縮されているならば、M D C T も位相のずれた余弦波を基底として用いているので、フーリエ変換との違いは位相のずれだけであり、M D C T 空間と周波数空間の間には性質のよい対応関係が期待できる。しかし最新の圧縮技術はオーディオデータの特性に応じて、乗じる窓関数の形状やその長さ（以下では窓長と呼ぶ）を変更させ音質の改善を達成している。このため M D C T のある周波数とフーリエ変換のある周波数を対応させるような単純な関係は得られず、計算式によって求めることはできないためテーブルに保存しておく必要がある。

【 0 0 2 2 】

図 2 に窓長および窓関数の具体例を図示する。本発明は、種々の圧縮データの規格に適応可能であるが、具体的に詳細に説明すべく、以下本発明の実施例では、

MPEG 2 の規格に基づき説明を行う。たとえばMPEG2 AAC(Advanced Audio Coding) では通常2048サンプルを窓長とする窓関数を乗じMDCTを行うが、音声之急激に変化する部分ではプリエコーと呼ばれる劣化を防ぐために256サンプルを窓長とし窓関数を乗じMDCTを行っている。2048サンプルを単位とする通常のフレームはONLY_LONG_SEQUENCEと呼ばれ1回のMDCTをほどこした結果である1024本のMDCT係数で記述され、256サンプルを単位とするフレームはEIGHT_SHORT_SEQUENCEと呼ばれ窓の半分ずつを重複させた256サンプル8回のMDCTの結果である128本のMDCT係数が8組で記述される。さらにこれらをつなぐためにLONG_START_SEQUENCEとLONG_STOP_SEQUENCEと呼ばれる左右非対称な窓関数も用いられる。

【0023】

図3に窓関数とMDCT係数列の関係について図示する。MPEG2 AACの場合には時間軸上のオーディオデータはこれらを用いて、たとえば図3の曲線のような順番で窓関数がかけられ、太線矢印のような順番でMDCT係数列が記述される。このような窓長の変化がある時には、フーリエ変換の基底は少数のMDCT係数に単純に変換されることはできない。

【0024】

従って、本発明の対応テーブルは、付加情報の埋め込みは窓関数に依存しないようにする。(付加情報埋め込みの際に加えられる信号は、圧縮状態をほどこし時間軸上に展開した際には窓関数に依存しない信号になっていること)。これにより、窓関数の形状や窓長に依存した埋め込み方法を使った場合、圧縮状態での埋め込み・検出は可能とするとともに、圧縮がほどこされた後にはどのような窓関数が使われていたのかを知ることができる。

次に、本発明の対応テーブルは、付加情報を埋め込むフレーム間で干渉がないように作成する。つまり付加情報の埋め込みはMDCTの窓を単位として行わない。時間軸上に展開された時には必ず一定のサンプル数に1ビットが埋め込まれるように埋め込みはなされなければならない。このサンプル数を1フレームと呼ぶ。MDCTは50%ずつ窓かけの対象を重複させるため、複数のフレームにまたがる窓が必ず存在する(図4のブロック3がこれに当たる)。単純にこのフレー

ムに埋め込みを行うと、複数のフレームにその影響が及んでしまう。逆に埋め込みを行わないと埋め込みが弱くなり検出成績が悪くなる。このフレームの前半と後半には異なった付加情報を表す信号を埋め込む。

対応テーブルが用いられるのは、付加情報の埋め込みの際にMDC T 係数から周波数成分を算出する時と、周波数空間で求めた埋め込み信号を再びMDC T 係数へと変換する時、そして検出の際には周波数空間での検出に相当する演算をMDC T 空間で行う時である。更新の際には検出と埋め込みを順に行うことになるので、前述のすべての変換が行われる。

【0 0 2 5】

[窓関数の長さが変わらない場合の対応テーブル作成方法]

まず窓長が一定である場合のテーブルの作成法とそれを用いた検出・埋め込み方法を説明する。後にこれらを複数の窓長へと拡張する。MDC T 係数は、時間軸上でNサンプルのオーディオデータに対して窓関数を乗算してMDC T を施した結果であるN/2本の係数ずつ1ブロックとして記述されているものとする（すなわち、一定の窓長をNサンプルとしている）。以下、この「ブロック」という用語では特記しない限りN/2本のMDC T 係数を表す。連続する2ブロックに対応する時間軸上オーディオデータは、50%すなわちN/2サンプルが重複している。

【0 0 2 6】

本発明が対象とするのはN/2の整数倍のサンプル数に対して1ビットという埋め込み率に限定される。ここでは1ビットの埋め込みをする時間軸上のサンプル数を $n \times N/2$ として、これを1フレームと呼ぶ。先に述べたような50%重複の性質に起因して、時間軸上で連続する2つのフレームにまたがるブロックも存在する。図4は $n=2$ の場合で時間軸上の2フレームと、それにMDC T 空間で対応する5ブロックの模式図である。図4中で下段は時間軸上のオーディオデータを、上段はMDC T 係数列を表し、楕円弧はMDC T の対象を表す。Block3はFrame 1とFrame 2にまたがるブロックである。

【0 0 2 7】

埋め込みはフレームごとに独立して行われるので、テーブルはフレーム単位で周波数成分とMDC T 係数の対応をとればよく、また逆に言えば隣接するフレ

ームに対する埋め込みは影響を及ぼしあってはいけない。そこで周期が $N/2 \times m$ であるフーリエ変換の各基底について以下の方法で求めたMDC T 係数列をもってテーブルを構成する。ここで m は $N/2$ 以下の整数である。図5は $n=2$ 、 $m=1$ の正弦波の場合の模式図である。

【0 0 2 8】

1フレームに関係するブロックは $n+1$ 個存在するが、このうち先頭と最後のブロックは前後のフレームにもまたがっている(図5中ではブロック1と3)。そこで振幅1.0で長さが1フレーム分の基底波形の前後に、値ゼロを持つサンプルを $N/2$ ずつつなぎあわせた波形を考える(図5中では太線部分がそれに当たる)。この波形の先頭から50%重複させながら N サンプルずつに対して窓関数を乗じ(図5中の楕円弧に対応する)、MDC Tをほどこせばこの波形のMDC T表現が得られる。逆にここで得られたMDC T係数列をIMDC Tすれば前後 $N/2$ サンプルずつはゼロ値となっている。

【0 0 2 9】

図6に隣接するフレームに付加情報を埋め込む例を図示する。図6のようにゼロ値のサンプルを補うことで、埋め込みの際に、隣接するフレームへの埋め込みを干渉させないことができる。検出および周波数成分の計算の時には前後のフレームに影響されない、そのフレームだけの検出結果や周波数成分を求めることができる。ゼロ値を補わない方法では埋め込みも検出も隣接するフレームと影響を及ぼしあってしまう。

【0 0 3 0】

テーブル作成の手順は以下の通りである。

ステップ1： まず周期 $N/2 \times n/k$ 、振幅1.0、長さ $N/2 \times n$ の余弦波を作成する。この余弦波は $N/2 \times n$ サンプルに対してフーリエ変換を行う時の k 番目の基底に当たる。

$$\begin{aligned} f(x) &= \cos(2\pi / (N/2 \times n/k) \times x) & (0 \leq x < N/2 \times n) \\ &= \cos(4k\pi / (N \times n) \times x) \end{aligned}$$

ステップ2： 波形の先頭と末尾に $N/2$ サンプルずつゼロ値のサンプルを補う(図5)。

$$\begin{aligned}
 g(y) &= 0 & (0 \leq y < N/2) \\
 &f(y - N/2) & (N/2 \leq y < N/2 \times (n+1)) \\
 &0 & (N/2 \times (n+1) \leq y < N/2 \times (n+2))
 \end{aligned}$$

ステップ 3 : $N/2 \times (b-1)$ 番目のサンプルから $N/2 \times (b+1)$ 番目のサンプルまでを取り出す。b は 1 から $n+1$ までの整数でありそのすべてについて以降の処理を行う。

$$h_b(z) = g(z + N/2 \times (b-1)) \quad (0 \leq z < N)$$

ステップ 4 : 窓をかける。

$$h_b(z) = h_b(z) \times \text{win}(z) \quad (0 \leq z < N, \text{win}(z) \text{ は窓関数})$$

ステップ 5 : MDCT を施し、結果として得られる $N/2$ 本の MDCT 係数をベクトル $V_{r, b, k}$ とする。

$$V_{r, b, k} = \text{MDCT}(h_b(z))$$

MDCT 変換は直交変換でありフーリエ変換の各基底は 1 次独立であるので、1 から $N/2$ までの値をとる k についての各 $V_{r, b, k}$ は直交している。

ステップ 6 : すべての (k, b) の組み合わせについて $V_{r, b, k}$ を求めた後に各行列 $T_{r, b}$ を構成する。

$$T_{r, b} = (V_{r, b, 1}, V_{r, b, 2}, V_{r, b, 3} \cdots V_{r, b, N/2})$$

同様の方法で正弦波について得たベクトルを $v_{i, b, k}$ 、行列を $T_{i, b}$ とする。その各列は大きさの 1 の正弦波を表す MDCT 係数列である。そしてブロック番号 b は 1 から $n+1$ までであるので、行列は $2 \times (n+1)$ 個となる。

【0 0 3 1】

周波数空間から MDCT 空間への変換

オーディオデータの周波数空間での表示を $R + jI$ とする。ここで j は虚数、 R はオーディオデータの実数成分を I は虚数成分を表す $N/2$ 次の実数ベクトルであり、その k 成分は $(N/2) \times n / k$ サンプルの周期を持つ基底に対応する。求める MDCT 係数列 M_b は、各周波数成分を別々に MDCT 空間へと変換した MDCT 係数列のベクトル和であるので、 $M_b = T_{r, b} + T_{i, b} I$ として計算できる。ここで b は 1 から $n+1$ までの整数で各ブロックに対応する。 M_1 と M_{n+1} は隣接するフレームにまたがるブロックの MDCT 係数列となっている。

【0 0 3 2】

MDC T空間から周波数空間への変換

各 $v_{i,b,k}$ 、 $v_{r,b,k}$ は直交してMDC T空間を張っているので、あるMDC T係数列 M_b を与えられた時にそれと各 $v_{r,b,k}$ 、 $v_{i,b,k}$ の内積をとれば M_b のその方向の成分を求めることができ、これがそれぞれそのまま周波数空間での実数成分と虚数成分を表す。1フレームに関する $(n+1)$ ブロックのMDC T係数列をまとめて処理して、そのフレームの周波数成分を求める式になっている。

【数 3】

$$R = \sum_{b=1}^{n+1} T_{r,b}^T M_b$$

$$I = \sum_{b=1}^{n+1} T_{i,b}^T M_b$$

【0 0 3 3】

〔窓関数がオーディオデータ中で変化する場合の対応テーブル作成方法〕

どのような窓関数が圧縮に用いられる可能性があるかは列挙されているものとする。またすべての窓長はそのうちの最大の窓長 N の約数であるとする。窓長が N/W サンプル (W は整数) のブロックでは、50%重複させながら N/W サンプルに対してMDC Tを W 回ほどこした結果として $N/(2W)$ 本のMDC T係数が W 組、合計で $N/2$ 本の係数が記述されているものとする。その W 回のうちの先頭のMDC Tはブロックのoffsetサンプル目から始まる N/W サンプルを変換するものとする。たとえばMPEG2 AACのEIGHT_SHORT_SEQUENCEの場合には $N=2048$ 、 $W=8$ 、offset=448であり、50%重複させながら256サンプルに対してMDC Tを8回ほどこした結果として128本のMDC T係数が8組、時間順に記述されている (図2 および、図3 参照)

テーブルの作成方法

窓長 N/W についてのテーブルは次のように作成される。。

ステップ1： 窓関数の長さが変わらない場合と同様。

ステップ2： 窓関数の長さが変わらない場合と同様。

ステップ3： w 個目の窓に相当する N/W サンプルを取り出す。 w は1から W までの整数値をとる。 b は1から $n+1$ までの整数値をとる。以降の処理は b と w のすべての組み合わせについてされなければならない。

$$h_{b,w}(z) = g(zN/2 \times (b-1) + N/2/W \times w + \text{offset}) \quad (0 \leq z < N/W)$$

ステップ4： 窓をかける。

$$h_{b,w}(z) = h_{b,w}(z) \times \text{win}(z) \quad (0 \leq z < N/W : \text{win}(z) \text{ は窓関数})$$

ステップ5： MDCTを施し、結果として得られる $N/(2W)$ 本のMDCT係数を $u_{r,b,k,w}$ に保存する。

$$u_{r,b,k,w} = \text{MDCT}(h_{b,w}(z))$$

ステップ6： $u_{r,b,k,w}$ を並べて $u_{r,b,k}$ とする

1から W までの値をとるすべての w について $u_{r,b,k,w}$ を求めたら、それらを縦に並べたベクトルが $u_{r,b,k}$ となる。

図7は $n=2$ 、 $b=2$ 、 $k=1$ 、 $W=8$ の場合、 $u_{r,2,1,w}$ が、この基底のどの部分をMDCTした係数列であるかを示している。

ステップ7： すべての (k,b) の組み合わせについて $u_{r,b,k}$ を求めた後に1から $N/2$ までの k について $u_{r,b,k}$ を横に並べて $T_{W,r,b}$ を構成する。

【0034】

各 $u_{r,b,k,w}$ は $N/(2w)$ 行1列のベクトルであるので、この行列は $N/2$ 行 $N/2$ 列の正方行列である。この各列は大きさの1の余弦波が b 番目に現れた窓長 N/W のブロックでどのようにMDCT係数列として表現されるかを表している。同様に正弦波についても行列 $T_{W,i,b}$ を求める。ブロック番号 b は1から $n+1$ までであるので、この窓長に対する行列は $2 \times (n+1)$ 個となる。さらに、窓長や窓関数の種類に応じてこのテーブルを作成する。

【0035】

・周波数空間からMDC T空間への変換

窓長が1種類の場合と異なるのは、圧縮されたオーディオデータからブロック情報を読み取ってブロックごとにどのような窓関数がいれたかに応じて異なった行列を用いる点である。それぞれのブロックごとに行列を変化させることで、どのような窓関数と窓長が使われていたとしてもそれに対応するようにMDC T係数列 M_b は調整され、これをIMDC Tして時間領域に変換した時に得られる波形、および、それをさらにフーリエ変換して周波数領域へ変換して得られる周波数成分は窓関数と窓長に依存しない。この M_b は、 $M_b = T_{w,r,b} R + T_{w,i,b} I$ として計算される。

【0 0 3 6】

・MDC T空間から周波数空間への変換

同様に $T_{r,b}$ の代わりに $T_{w,r,b}$ を用いれば周波数空間への変換も同様に行うことができる。窓関数と窓長に対応して行列を変化させることで、窓関数と窓長に依存しない真の周波数成分が求められる。

【数 4】

$$R = \sum_{b=1}^{n+1} T_{w,r,b}^T M_b$$

$$I = \sum_{b=1}^{n+1} T_{w,i,b}^T M_b$$

【0 0 3 7】

[テーブルに必要な記憶容量の縮小を行う方法]

行列は $(N/2) \times (N/2)$ の大きさを持つので、この方法で作成されるテーブルは一つの窓関数について $2 \times (n+1) \times (N/2) \times (N/2) = (n+1) \times N^2 / 2$ 個のMDC T係数（浮動小数点数）で構成されることになる。しかしこのテーブルの

内容は冗長性が高いので実際に必要な記憶容量は大幅に縮小することができる。

【0038】

方法1：基底の周期性を利用する方法

まず1つの方法として基底の周期性を利用することができる。この方法では V_r , b , k のうち幾つかがまったく同じものであることに注目しその部分を省く。
 m を整数としたとき、 $N/2 \times m$ サンプル先の余弦波は

$$\begin{aligned} f(x+N/2 \times m) &= \cos(4k\pi / (N \times n) \times (x+N/2 \times m)) \\ &= \cos(4k\pi / (N \times n) \times x + 4k\pi / (N \times n) \times N/2 \times m) \\ &= \cos(4k\pi / (N \times n) \times x + 2\pi k \times m/n) \end{aligned}$$

なので、 $[a] \ (k \times m)/n$ が整数である場合

$$\begin{aligned} f(x+N/2 \times m) &= f(x) \quad (0 \leq x \leq N/2 \times (n-m) \text{ の範囲に限る }) \\ g(y+N/2 \times m) &= g(y) \quad (N/2 \leq y \leq N/2 \times (n-m+1) \text{ の範囲に限る }) \end{aligned}$$

であるので

$$h_{b+m}(z) = h_b(z) \quad (2 \leq b \leq n-m \text{ の範囲に限る })$$

となって

$$V_{r, b+m, k} = V_{r, b, k} \quad (2 \leq b \leq n-m \text{ の範囲に限る })$$

となる。範囲の制限は $f(x)$ の定義域を理由とする。

$[b] \ (k \times m)/n$ が整数/2で表現できる既約分数である場合

$$f(x+N/2 \times m) = -f(x)$$

であり

$$h_{b+m}(z) = -h_b(z)$$

であることから

$$V_{r, b+m, k} = -V_{r, b, k}$$

となる。範囲の制限は $[a]$ と同様。

$[c] \ (k \times m)/n$ が $(4 \times \text{整数} + 1)/4$ で表現できる既約分数である場合

$$\begin{aligned} f(x+N/2 \times m) &= \cos(4k\pi / (N \times n) \times x + \pi (\text{偶数} + 1/2)) \\ &= -\sin(4k\pi / (N \times n) \times x) \end{aligned}$$

であるので

$$V_{r, b+m, k} = -V_{i, b, k}$$

[d] $(k \times m)/n$ が $(4 \times \text{整数} + 3)/4$ で表現できる場合

$$\begin{aligned} f(x + N/2 \times m) &= \cos(4k\pi / (N \times n) \times x + \pi(\text{奇数} + 1/2)) \\ &= \sin(4k\pi / (N \times n) \times x) \end{aligned}$$

であるので

$$V_{r, b+m, k} = V_{i, b, k}$$

となる。範囲の制限は [a] と同様。

【0 0 3 9】

よって [a] から [d] のいずれかの条件を満たす $V_{r, b+m, k}$ は他のベクトルで代用ができる。 $V_{i, b, k}$ についても同様である。よって、行列 $T_{r, b}$ と行列 $T_{i, b}$ を行列としてそのまま記憶しておくのではなく、以下の最小の構成要素を記憶しておくので十分である。最小の構成要素とは以下の通りである。

【0 0 4 0】

- ・ [a] ~ [d] の条件を満たさないベクトル $V_{r, b, k}$ および $V_{i, b, k}$
- ・ 行列 $T_{r, b}$ と $T_{i, b}$ の各列としてどのベクトルを正負どちらの符号で使うかの情報

【0 0 4 1】

MDC T 空間と周波数空間の間の変換を実際にやる際には、行列 $T_{r, b}$ や行列 $T_{i, b}$ の各列の代わりに $V_{r, b, k}$ および $V_{i, b, k}$ を用いて、行列演算と等価な演算を行うことができる。周波数空間から MDC T 空間への変換は次式となる。

【数 5】

$$\begin{aligned} M_b &= T_{r, b} R + T_{i, b} I \\ &= \sum_{k=1}^{N/2} (R_k V_{r, b, k} + I_k V_{i, b, k}) \end{aligned}$$

ベクトルを共通化したところでは適当に他のベクトルを用いる。MDC T 空間から周波数空間への変換は各周波数成分ごとに、以下の内積を求めることで行う。

この式は行列 $T_{r, b}$ や行列 $T_{i, b}$ を使う場合の式を各成分に分解した式となる。

【数 6】

$$R_k = V_{r,b,k} \cdot M_b$$

$$I_k = V_{i,b,k} \cdot M_b$$

必要な記憶容量がこの共通化によって減る程度は n に依存する。たとえば $n=3$ の時は[a]しか成立しえないので8.3%しか減らないが、 $n=4$ の時は40%が減る。

窓関数が増える場合も hb, w に、窓関数が一通りしかない場合と同様の関係があるので上述の共通化はそのまま適用できて、同様の条件が満たされた時に次式となる。

【数 7】

$$[a] \quad u_{r,b+m,k} = u_{r,b,k}$$

$$[b] \quad u_{r,b+m,k} = -u_{r,b,k}$$

$$[c] \quad u_{r,b+m,k} = -u_{i,b,k}$$

$$[d] \quad u_{r,b+m,k} = u_{i,b,k}$$

【0 0 4 2】

方法 2 : 基底を前後に分解する方法

さらにMDCTの線形性を利用して、フーリエ変換の基底を部分々に分解し、それを変換したMDCT係数列をテーブルにすれば前述の方法 1 の適用範囲を

広げることができる。変換の際にはテーブルに記憶されたMDC T係数列のベクトル和で基底を表現する。図8に基底に分解例を図示する。

まず波形(図8左端、太線)を各ブロックごとに前半のN/2サンプルと後半のN/2サンプルに分け、前半をMDC Tする際には後半にゼロ値の波形をN/2サンプル補ってMDC Tを行い(図8中央)、後半をMDC Tする際には前半にゼロ値の波形をN/2サンプル補ってMDC Tを行う(図8右端)。ここでは波形の前半(後半)をMDC Tして得られたMDC T係数列をベクトル $V_{fore, r, b, k}$ ($V_{back, r, b, k}$) で表すことにする。MDC Tには線形性があるので元の波形のMDC T係数列 $V_{r, b, k}$ は $V_{fore, r, b, k}$ と $V_{back, r, b, k}$ のベクトル和に等しい。

【0 0 4 3】

このように分解すると方法1では $V_{r, b, k}$ を共通化できなかった部分でも $V_{fore, r, b, k}$ や $V_{back, r, b, k}$ を共通化できるようになる。たとえば図5においてBlock1は $b=1$ なので前述の方法1は適用不可能だった。しかし各ブロックを前後に分解して考えるとBlock1のMDC T係数列 $V_{back, r, 1, k}$ とBlock2のMDC T係数列 $V_{back, r, 2, k}$ は正負が反転するのみなので一方の記憶を省けることがわかる。Block2の $V_{fore, r, 2, k}$ とBlock3の $V_{fore, r, 3, k}$ も同様であり、そしてBlock1の $V_{fore, r, 1, k}$ とBlock3の $V_{back, r, 3, k}$ は常にゼロベクトルになる。

【0 0 4 4】

この方法を使ったテーブル作成の手順は以下の通りである。

ステップ1： 基底を前後に分解しない場合と同様。

ステップ2： 基底を前後に分解しない場合と同様。

ステップ3： まず fore 係数列の作成。N/2×(b-1)番目からN/2×b番目を取り出しその後にゼロ値のN/2サンプルを補う。

$$h_{fore, b}(z) = \begin{cases} g(z+N/2 \times (b-1)) & (0 \leq z < N/2) \\ 0 & (N/2 \leq z < N) \end{cases}$$

ステップ4： 窓をかける。

$$h_{fore, b}(z) = h_{fore, b}(z) \times \text{win}(z) \quad (0 \leq z < N, \text{win}(z) \text{ は窓関数})$$

ステップ5： MDC Tを施し、結果として得られるN/2本のMDC T係数をベク

トル $V_{\text{fore}, r, b, k}$ とする。

$$V_{\text{fore}, r, b, k} = \text{MDCT}(h_{\text{fore}, b}(z))$$

ステップ 6 : 次に back 係数列の作成。 $N/2 \times b$ 番目から $N/2 \times (b+1)$ 番目を取り出しその前にゼロ値の $N/2$ サンプルを補う。

$$h_{\text{back}, b}(z) = 0 \quad (0 \leq z < N/2)$$

$$g(z + N/2 \times (b-1)) \quad (N/2 \leq z < N)$$

ステップ 7 : 窓をかける。

$$h_{\text{back}, b}(z) = h_{\text{back}, b}(z) \times \text{win}(z) \quad (0 \leq z < N, \text{win}(z) \text{ は窓関数})$$

ステップ 8 : MDCT を施し、結果として得られる $N/2$ 本の MDCT 係数をベクトル $V_{\text{back}, r, b, k}$ とする。

$$V_{\text{back}, r, b, k} = \text{MDCT}(h_{\text{back}, b}(z))$$

ステップ 9 : すべての (k, b) の組み合わせについて $V_{\text{fore}, r, b, k}$ と $V_{\text{back}, r, b, k}$ を求めた後に各行列 $T_{\text{fore}, r, b}$ と $T_{\text{back}, r, b}$ を構成する。

$$T_{\text{fore}, r, b} = (V_{\text{fore}, r, b, 1}, V_{\text{fore}, r, b, 2} \cdots V_{\text{fore}, r, b, N/2})$$

$$T_{\text{back}, r, b} = (V_{\text{back}, r, b, 1}, V_{\text{back}, r, b, 2} \cdots V_{\text{back}, r, b, N/2})$$

【0 0 4 5】

MDCT の線形性から

$$V_{r, b, k} = V_{\text{fore}, r, b, k} + V_{\text{back}, r, b, k}$$

であり、

$$T_{r, b} = T_{\text{fore}, r, b} + T_{\text{back}, r, b}$$

である。この性質を利用し MDCT 空間と周波数空間の間の変換では $T_{r, b}$ を用いるのと等価な操作を $T_{\text{fore}, r, b}$ と $T_{\text{back}, r, b}$ を用いて行えばよい。

【0 0 4 6】

ここで、これらの定義の下で基底の周期性を利用すると

[a] $(k \times m)/n$ が整数である場合

$b+m=n+1$ という条件においても

$$h_{\text{fore}, n+1}(z) == h_{\text{fore}, b}(z)$$

が成立する。これは $h_{\text{fore}, b}(z)$ の後半がゼロ値であるからである。よって下式の適用範囲が広くなり

$$h_{\text{fore}, b+m}(z) == h_{\text{fore}, b}(z) \quad (2 \leq b \leq n-m+1 \text{ の範囲に限る })$$

であり

$$V_{\text{fore}, r, b+m, k} == V_{\text{fore}, r, b, k} \quad (2 \leq b \leq n-m+1 \text{ の範囲に限る })$$

となり共通化される部分が多くなる。 $V_{\text{back}, r, b, k}$ では $b=1$ という条件でも
$$h_{\text{back}, m+1}(z) == h_{\text{back}, 1}(z)$$

が成立する。これは $h_{\text{back}, 1}(z)$ の前半がゼロ値であるからである。よって下式の適用範囲が広くなり

$$h_{\text{back}, b+m}(z) == h_{\text{back}, b}(z) \quad (1 \leq b \leq n-m \text{ の範囲に限る })$$

であるため

$$V_{\text{back}, r, b+m, k} == V_{\text{back}, r, b, k} \quad (1 \leq b \leq n-m+1 \text{ の範囲に限る })$$

となり共通化される部分が多くなる。 $[b] [c] [d]$ についても範囲の制限はこれと同条件になる。

【0 0 4 7】

方法3：近似する方法

テーブルを縮小する最後の方法は近似である。フーリエ変換の基底波形1本に対応するMDCT係数列のうち、ある程度より小さい値を持つMDCT係数はゼロに近似しても実用上問題はおきない。この近似に用いる閾値には変換の精度と記憶容量のトレードオフによって適当な値を選んで決める。そしてゼロと近似した部分は行列演算を行わないように各システムを設計することで計算時間も短縮することができる。

さらに、値の大きい係数も含めてすべての係数を有理数に近似し量子化してしまうことで浮動小数点数ではなく整数として記憶し容量を節約することもできる。

【0 0 4 8】

〔対応テーブル作成器〕

テーブル作成は、基本的に、窓に関する情報を入力として受け取り、テーブルを作成し出力することからなる。上記の対応テーブル作成方法と同様に、窓に関する情報とは、フレーム長 N 、フレームに対するブロックの長さを表す n 、先頭の窓のオフセット offset 、窓関数、窓長を規定する W である。テーブルは、基本的に

、対象の音声圧縮技術が使う窓の種類の数だけ作る。

【 0 0 4 9 】

〔付加情報埋め込みシステム〕

図 9 に本発明の付加情報埋め込みシステムのブロック構成図を示す。M D C T 係数復元部 (210) は、入力データである圧縮音声データから、音声の M D C T 係数列と窓情報とその他の情報を復元する。これらの情報は、入力データである圧縮音声データ内に指定された、ハフマン符号の復号、逆量子化、予測方法を用いて取り出（復元）される。次に M D C T / D F T 変換部 (230) は、M D C T 係数復元部 (210) において復元された音声の M D C T 係数列と窓情報を受け取り、テーブル (900) を用いて周波数成分に変換する。そして周波数空間埋め込み部 (250) は、M D C T / D F T 変換部 (230) において変換の結果得られた周波数成分に、付加情報の埋め込みを行う。

D F T / M D C T 変換部 (240) は、周波数空間埋め込み部 (250) で埋め込みを行われた周波数成分を、M D C T 係数復元部 (210) において取り出しておいた窓情報に従って、テーブル (900) を用いて M D C T 係数列へと変換する。最後に M D C T 係数圧縮部 (220) が、D F T / M D C T 変換部 (240) で得られた M D C T 係数列を、M D C T 係数復元部 (210) において取り出しておいた窓情報とその他の情報と併せて圧縮し、圧縮音声データを作成する。圧縮の際には窓情報とその他の情報が指示する予測方法、逆量子化、ハフマン符号化を用いて圧縮する。このように構成することにより、付加情報の埋め込みは周波数成分の操作に対応するように行われているため、圧縮がほどこれた後でも既存の周波数空間検出方法で検出を行うことができる。

【 0 0 5 0 】

〔付加情報検出システム〕

図 1 0 に本発明の付加情報検出システムのブロック構成図を示す。M D C T 係数復元部 (210) は、入力データである圧縮音声データから、音声の M D C T 係数列と窓情報とその他の情報を復元する。これらの情報は、入力データである圧縮音声データに指定された、ハフマン符号の復号、逆量子化、予測方法を用いて取り出（復元）される。次に M D C T / D F T 変換部 (230) は、M D C T 係数復元部 (2

10)において復元された音声のMDCT係数列と窓情報を受け取り、テーブル(900)を用いて周波数成分に変換する。最後に、周波数空間検出部は、MDCT/DFT変換部(230)において周波数成分に変換された情報から、埋め込まれた付加情報を検出し、これを出力する。

を、MDCT空間上で行う。

【0051】

[付加情報更新システム]

図11に本発明の付加情報更新システムのブロック構成図を示す。

MDCT係数復元部(210)は、入力データである圧縮音声データから、音声のMDCT係数列と窓情報とその他の情報を復元する。これらの情報は、入力データである圧縮音声データ内に指定された、ハフマン符号の復号、逆量子化、予測方法を用いて取り出(復元)される。

次にMDCT/DFT変換部(230)は、MDCT係数復元部(210)において復元された音声のMDCT係数列と窓情報を受け取り、テーブル(900)を用いて周波数成分に変換する。

周波数空間更新部(410)は、MDCT/DFT変換部(230)において得られた周波数成分の中に付加情報が埋め込まれているかどうかをまず判定する。埋め込まれているなら、その内容を変更する必要があるかをさらに判定する。その必要がある場合のみ、付加情報の更新を周波数成分に対して行う。(更新器の利用者にわかるように、それぞれの判定の結果を出力してもよい。)

DFT/MDCT変換部(240)は、周波数空間更新部(250)において付加情報の更新を行われた周波数成分を、MDCT係数復元部(210)において取り出しておいた窓情報に従って、テーブル(900)を用いてMDCT係数列へと変換する。

最後にMDCT係数圧縮部(220)が、DFT/MDCT変換部(240)で得られたMDCT係数列を、MDCT係数復元部(210)において取り出しておいた窓情報とその他の情報と併せて圧縮し、圧縮音声データを作成する。圧縮の際には窓情報とその他の情報が指示する予測方法、逆量子化、ハフマン符号化を用いて圧縮する。

【0052】

[一般的なハードウェア構成例]

本発明にかかる装置、システムは、通常のコンピュータのハードウェアを用いることにより実施可能である。図 1 2 に一般的なパーソナルコンピュータのハードウェア構成例を示す。システム 1 0 0 は、中央処理装置 (CPU) 1 とメモリ 4 とを含んでいる。CPU 1 とメモリ 4 は、バス 2 を介して、補助記憶装置としてのハードディスク装置 1 3 (または CD-ROM 2 6、DVD 3 2 等の記憶媒体駆動装置) と IDE コントローラ 2 5 を介して接続してある。同様に CPU 1 とメモリ 4 は、バス 2 を介して、補助記憶装置としてのハードディスク装置 3 0 (または MO 2 8、CD-ROM 2 9、DVD 3 1 等の記憶媒体駆動装置) と SCSI コントローラ 2 7 を介して接続してある。フロッピーディスク装置 2 0 はフロッピーディスクコントローラ 1 9 を介してバス 2 へ接続されている。

【0 0 5 3】

フロッピーディスク装置 2 0 には、フロッピーディスクが挿入され、このフロッピーディスク等やハードディスク装置 1 3 (または CD-ROM 2 6、DVD 3 2 等の記憶媒体)、ROM 1 4 には、オペレーティングシステムと協働して CPU 等に命令を与え、本発明を実施するためのコンピュータプログラム、ブラウザプログラム、オペレーティングシステムのコード若しくはデータを記録することができ、メモリ 4 にロードされることによって実行される。これらコンピュータ・プログラムのコードは圧縮し、または、複数に分割して、複数の記録媒体に記録することもできる。該プログラムを、ディスクセットなどの記録媒体に記録し、該ディスクセットを他のコンピュータで動作させることも可能である。

【0 0 5 4】

システム 1 0 0 は更に、ユーザ・インターフェース・ハードウェアを備え、入力をするためのポインティング・デバイス (マウス、ジョイスティック等) 7 またはキーボード 6 や、ディスプレイ 1 2 を有することができる。また、パラレルポート 1 6 を介してプリンタを接続することや、シリアルポート 1 5 を介してモデムを接続することが可能である。このシステム 1 0 0 は、シリアルポート 1 5 およびモデムまたは通信アダプタ 1 8 (イーサネットやトークンリング・カード) 等を介してネットワークに接続し、他のコンピュータ、サーバ等と通信を行うこ

とができる。またシリアルポート 1 5 若しくはパラレルポート 1 6 に、遠隔送受信機器を接続して、赤外線若しくは電波によりデータの送受信を行ってもよい。

【0 0 5 5】

スピーカ 2 3 は、オーディオ・コントローラ 2 1 によって D/A（デジタル／アナログ変換）変換されたサウンド、音声信号を、アンプ 2 2 を介して受領し、サウンド、音声として出力する。また、オーディオ・コントローラ 2 1 は、マイクロフォン 2 4 から受領した音声情報を A/D（アナログ／デジタル）変換し、システム外部の音声情報をシステムにとり込むことを可能にしている。音声をマイクロフォン 2 4 から入力し、これに基づき本発明にかかる圧縮データを作成してもよい。

上記ハードウェア構成は、通常のパーソナルコンピュータ（PC）のほか、ワークステーション、ノートブック PC、パームトップ PC、ネットワークコンピュータ、コンピュータを内蔵したテレビ等の各種家電製品、通信機能を有するゲーム機、電話、FAX、携帯電話、PHS、電子手帳、等を含む通信機能有する通信端末、または、これらの組合せによって実施可能であることを容易に理解できるであろう。ただし、これらの構成要素は例示であり、その全ての構成要素が本発明の実施に必要な必須の構成要素となるわけではないことに留意されたい。

【0 0 5 6】

【発明の効果】

本発明により、圧縮されたデジタル・オーディオデータに対する、付加情報の、埋め込み、検出、もしくは更新を圧縮された状態のまま直接行う方法およびシステムが提供される。さらに本発明の方法により、圧縮状態のオーディオデータに埋め込まれた付加情報は圧縮が解凍された後にも従来の電子透かし技術によって検出することができる。

【図面の簡単な説明】

【図 1】

圧縮オーディオデータに付加情報を直接埋め込む装置のブロック図である。

【図 2】

窓長および窓関数の具体例である。

【図 3】

窓関数とMDC T 係数列の関係を示す図である。

【図 4】

時間軸上のフレームと対応するMDC T 空間のブロックを示す図である。

【図 5】

正弦波の模式図である。

【図 6】

隣接するフレームに付加情報を埋め込む例である。

【図 7】

基底のどの部分をMDC T した係数列であるかを示す図である。

【図 8】

基底の分解例である。

【図 9】

本発明の付加情報埋め込みシステムのブロック構成図である。

【図 1 0】

本発明の付加情報検出システムのブロック構成図である。

【図 1 1】

本発明の付加情報更新システムのブロック構成図である。

【図 1 2】

一般的なコンピュータのハードウェア構成例である。

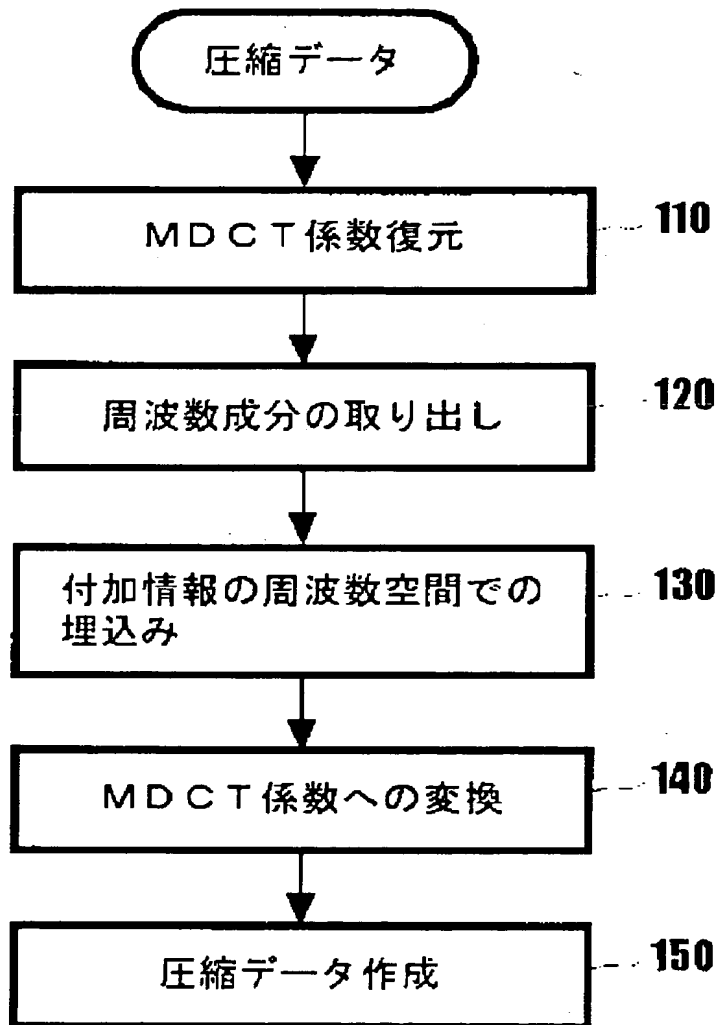
【符号の説明】

- 1 . . . CPU
- 2 . . . バス
- 4 . . . メモリ
- 5 . . . キーボード・マウス・コントローラ
- 6 . . . キーボード
- 7 . . . ポインティングデバイス
- 8 . . . ディスプレイ・アダプタ・カード
- 9 . . . ビデオメモリ

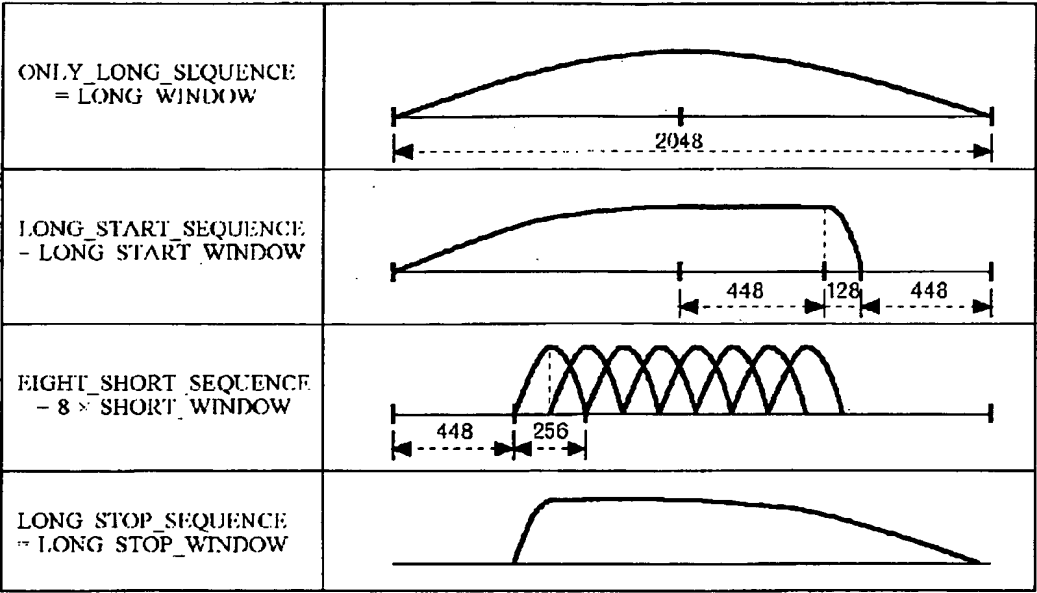
- 1 0 . . . D A C / L C D C
- 1 1 . . . 表示装置
- 1 2 . . . C R Tディスプレイ
- 1 3 . . . ハードディスク装置
- 1 4 . . . R O M
- 1 5 . . . シリアルポート
- 1 6 . . . パラレルポート
- 1 7 . . . タイマ
- 1 8 . . . 通信アダプタ
- 1 9 . . . フロッピーディスクコントローラ
- 2 0 . . . フロッピーディスク装置
- 2 1 . . . オーディオ・コントローラ
- 2 2 . . . アンプ
- 2 3 . . . スピーカ
- 2 4 . . . マイクロフォン
- 2 5 . . . I D Eコントローラ
- 2 6 . . . C D - R O M
- 2 7 . . . S C S Iコントローラ
- 2 8 . . . M O
- 2 9 . . . C D - R O M
- 3 0 . . . ハードディスク装置
- 3 1 . . . D V D
- 3 2 . . . D V D
- 1 0 0 . . . システム

【書類名】 図面

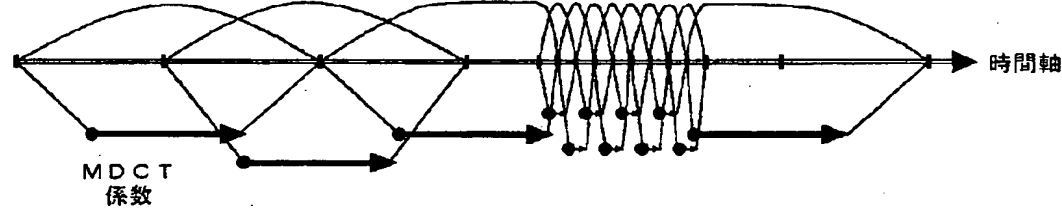
【図 1】



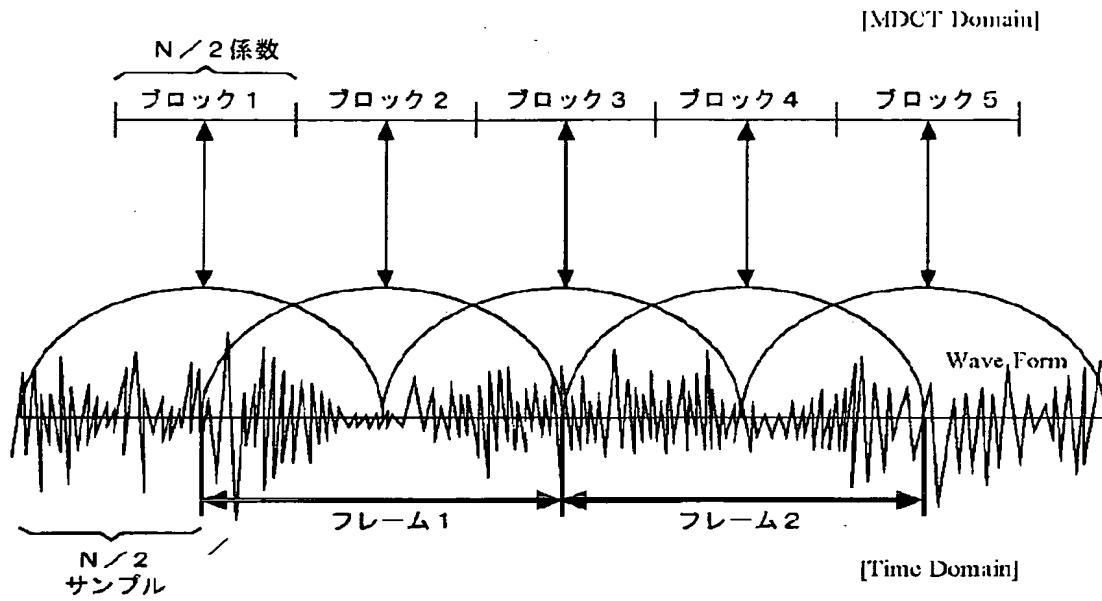
【図 2】



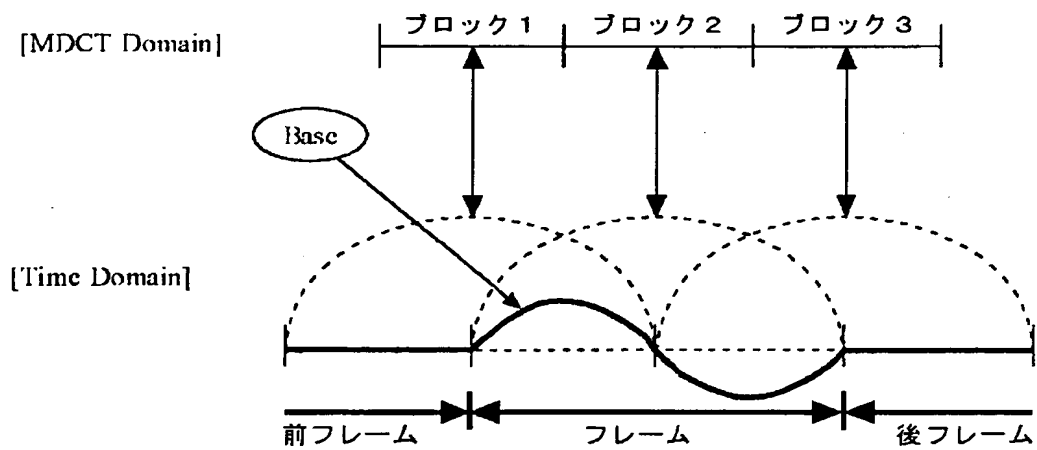
【図 3】



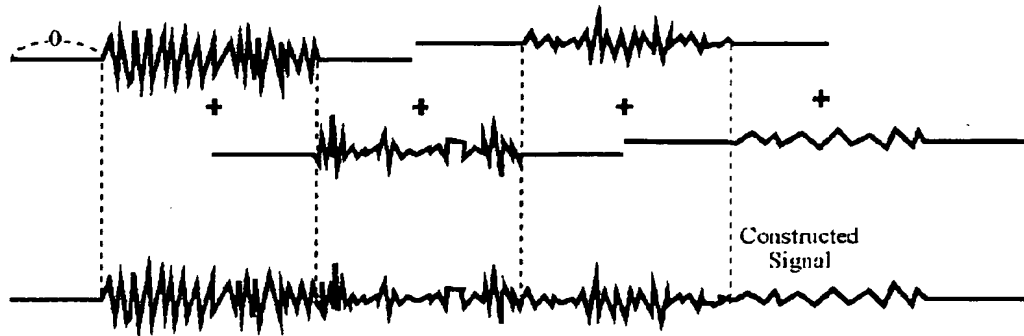
【図 4】



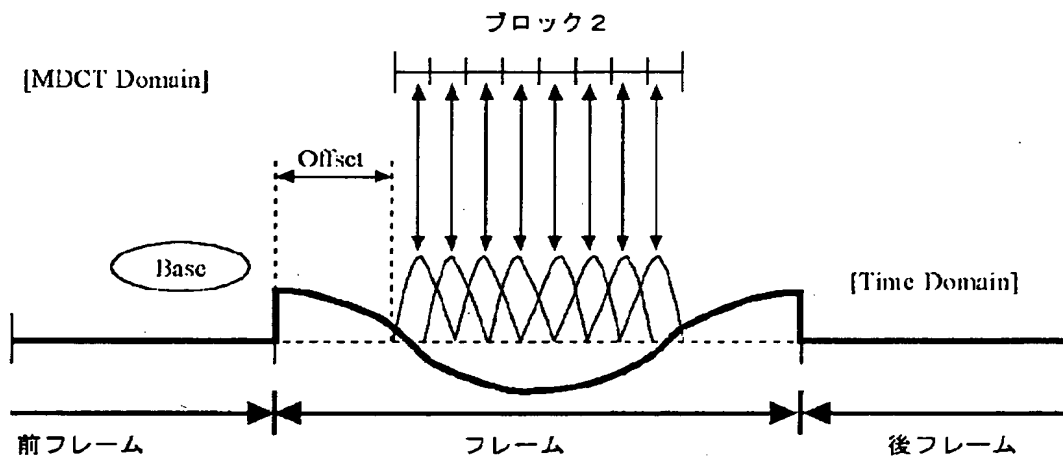
【図 5】



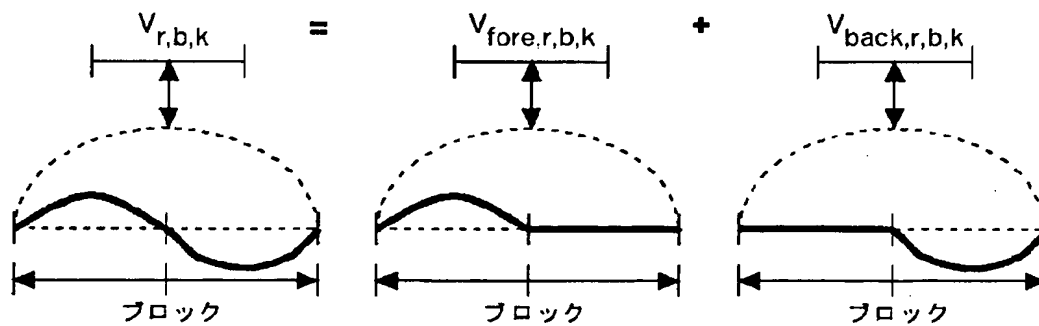
【図 6】



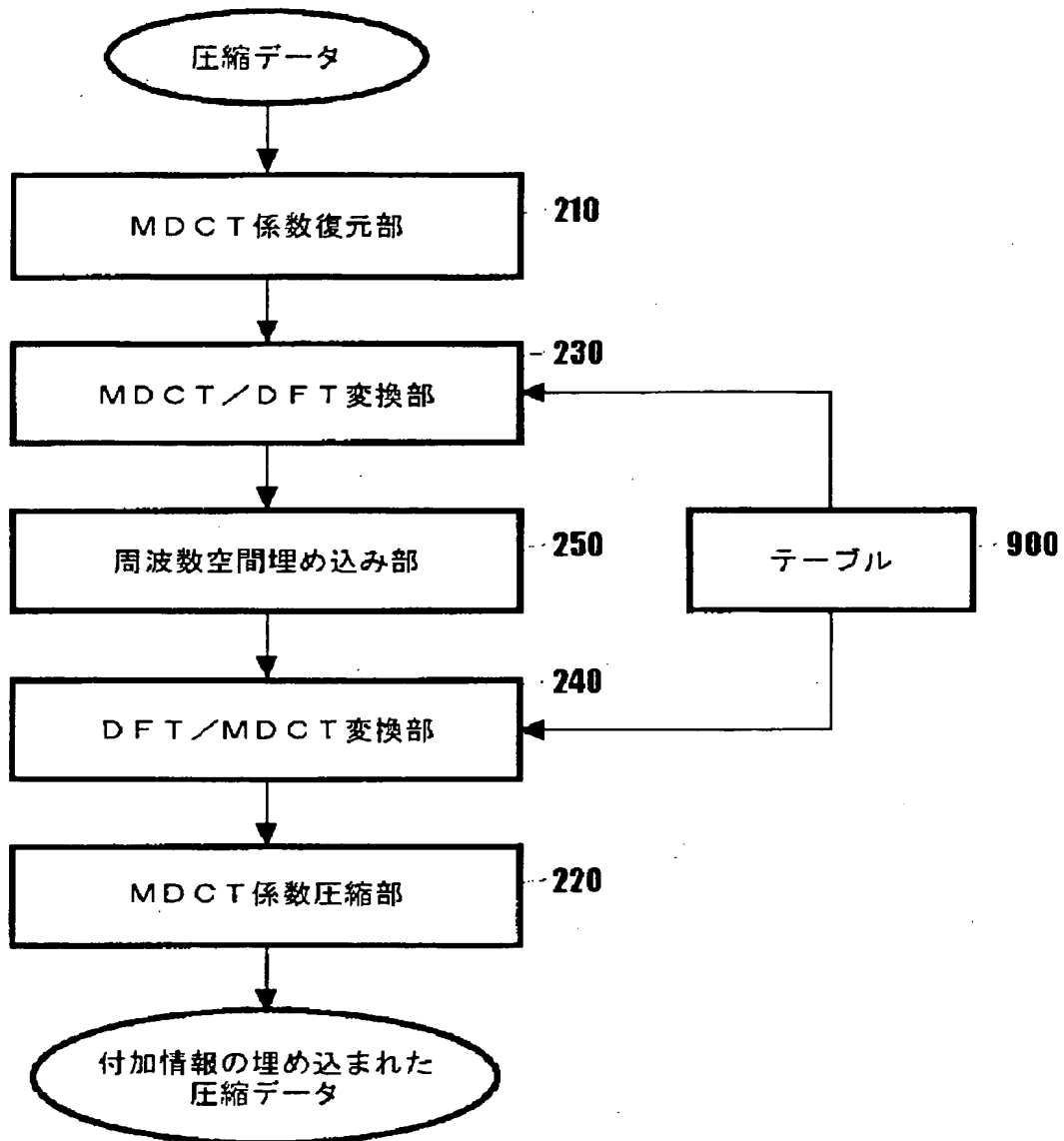
【図 7】



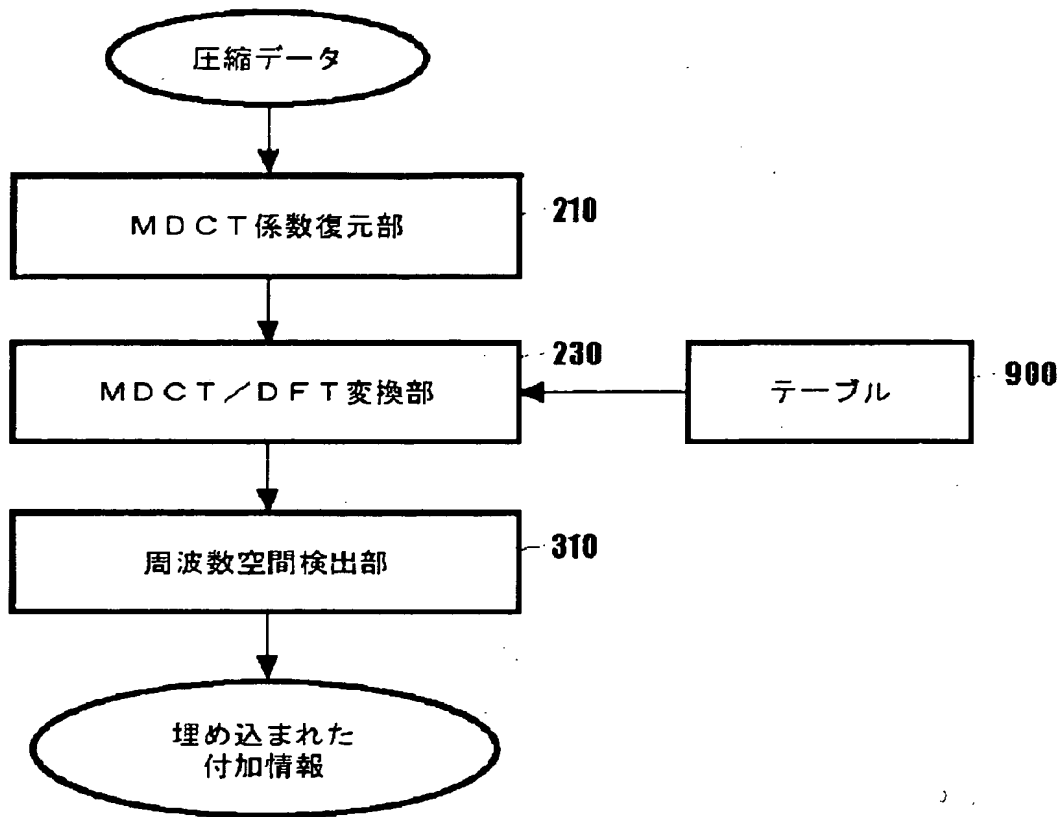
【図 8】



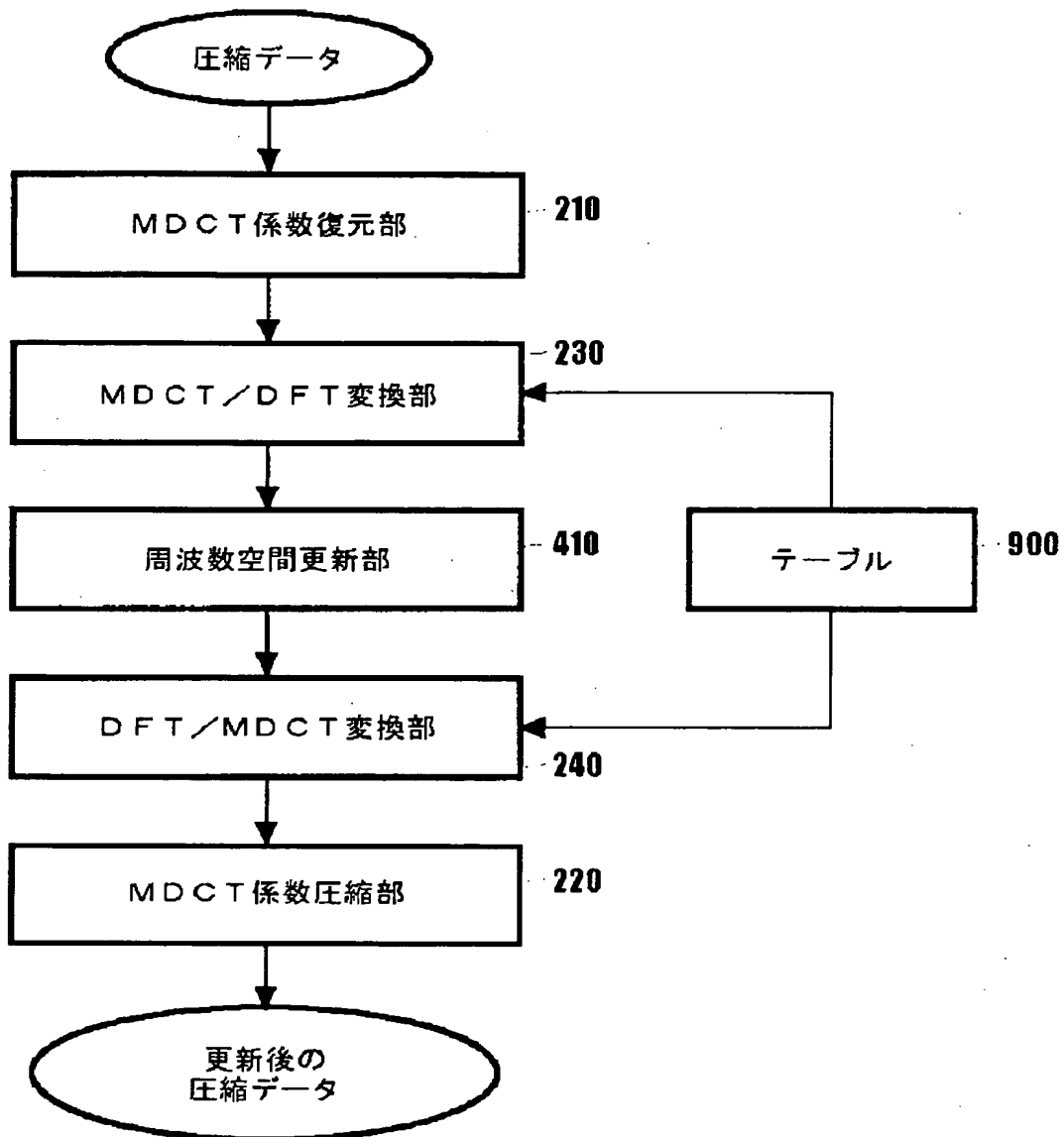
【図 9】



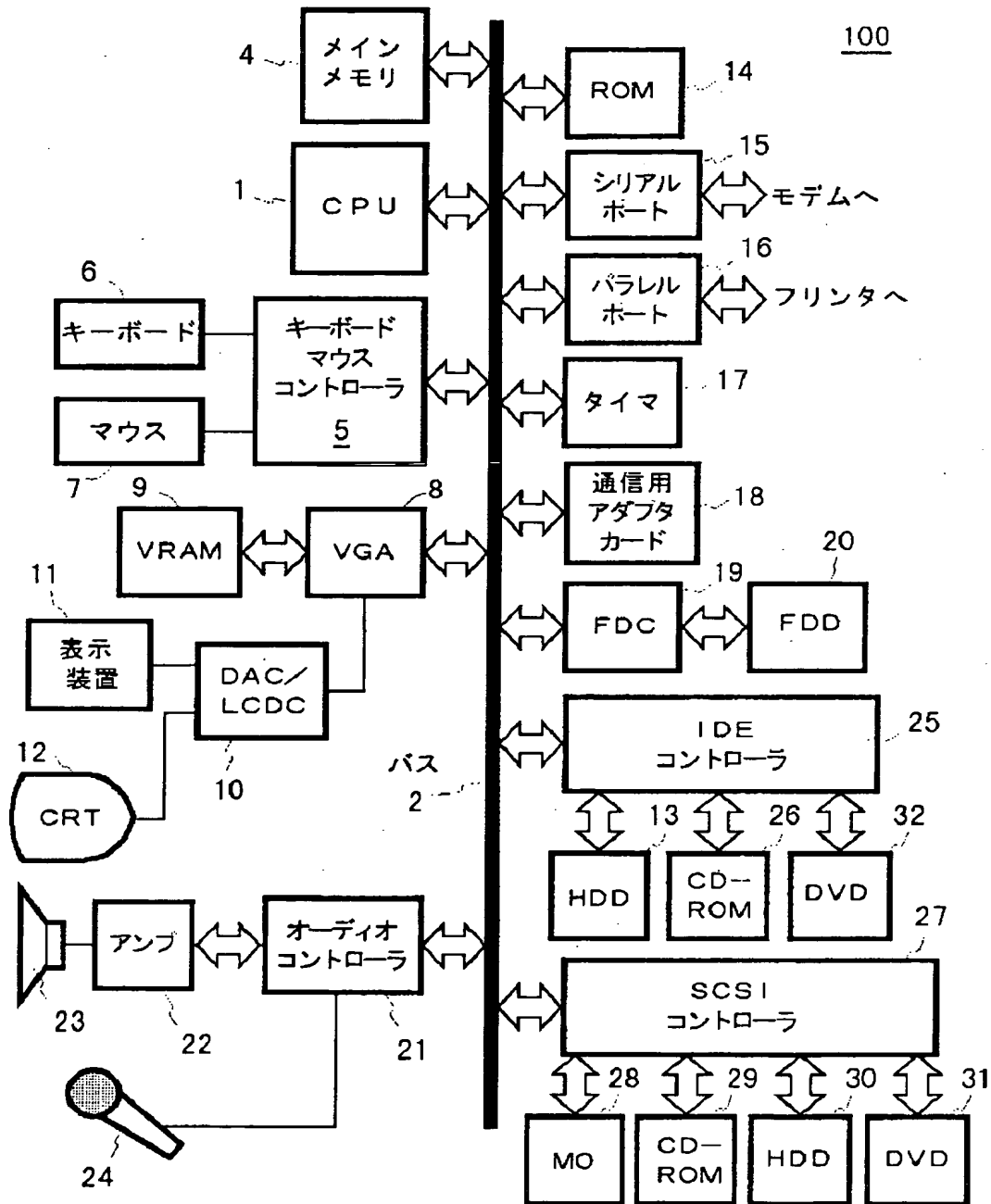
【図 1 0】



【図 1 1】



【図 1 2】



【書類名】 要約書

【要約】

【課題】

圧縮状態にあるデジタルオーディオデータ内の情報を直接操作する方法およびシステムを提供することである。

【解決手段】

圧縮オーディオデータに付加情報を埋め込むシステムは、

(1) 圧縮オーディオデータからMDC T (Modified Discrete Cosine Transform) 係数を復元する手段と、

(2) 復元された前記MDC T係数を用いて、オーディオデータの周波数成分を求める手段と、

(3) 求めた前記周波数成分に対して、付加情報を周波数空間で埋め込む手段と、

(4) 前記付加情報の埋め込まれた周波数成分をMDC T係数に変換する手段と

(5) 付加情報の埋め込まれた前記MDC T係数から圧縮オーディオデータを作成する手段、

を有する。

【選択図】 図 1

認定・付加情報

特許出願の番号	平成 1 1 年 特許願 第 3 6 4 6 2 7 号
受付番号	5 9 9 0 1 2 5 4 1 3 1
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 2 年 1 月 4 日

<認定情報・付加情報>

【提出日】	平成11年12月22日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 1990年10月24日
[変更理由] 新規登録
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション

2. 変更年月日 2000年 5月16日
[変更理由] 名称変更
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション